

Уточнение регрессионной многофакторной модели уровня воды в реке Ия (Восточная Сибирь)

С.И. Носков, Ю.Д. Аксенов, Ю. М. Сапожников

Иркутский государственный университет путей сообщения

Аннотация: В работе приведена уточненная регрессионная модель динамики уровня воды в сибирской реке Ия, включающая в правую часть шесть природных факторов (количество дней с осадками в горах Саян, средние дневная и ночная температуры за месяц, количество осадков, глубина снежного покрова, среднее атмосферное давление за месяц) с учетом запаздывания, а также специальным образом сформированную сезонную переменную. На высокую адекватность модели указывают значения критериев множественной детерминации, Фишера, средней относительной ошибки аппроксимации. Построенная модель может быть эффективно использована для решения широкого круга прогнозных задач.

Ключевые слова: регрессионная модель, уровень воды в реке, лаг запаздывания, сезонная переменная, прогноз.

Введение

При исследовании закономерностей функционирования гидрологических объектов, в том числе, уровня воды в реках, активно используются методы математического, в частности, регрессионного моделирования. Так, в работе [1] с использованием метода географически взвешенной регрессии производится оценка общего состояния качества воды в реках Индии. В [2] предлагается метод одновременной количественной оценки и анализа влияния предикторов климата на водоотдачу в бассейне реки Янцзы (Китай). В [3] предлагается новый подход к статистическому моделированию для ежечасного прогноза уровня реки на гидропосте в условиях потенциального риска наводнений при среднесрочном горизонте прогнозирования (около трех дней). С этой целью вводится модель регрессии переключения, которая учитывает последовательную корреляционную структуру рядов часовых уровней и изменяющуюся временную задержку между ними. Статья [4] посвящена применению многомерной линейной регрессии и логистической регрессии с несколькими пороговыми значениями для оценки загрязнения в водоразделе реки Тамсуи (северный Тайвань) с

использованием вспомогательных данных об окружающей среде. В [5] с использованием квантильной модели эмпирически измеряется предельное влияние ВВП на душу населения, уровня технологических инноваций, промышленной структуры, открытости, финансовой децентрализации и уровня урбанизации на сброс сточных вод в 11 провинциях вдоль экономического пояса реки Янцзы с 2008 по 2018 годы. Результаты моделирования показывают, что правительству необходимо срочно улучшить политику, например, произвести модернизацию экономических структур, предотвращение чрезмерной урбанизации, ускорение технологических инноваций, внедрение экологически безопасных иностранных инвестиций и предоставление большего вознаграждения лучшим практикам экологического управления, чтобы способствовать достижению так называемого зеленого экологического развития. В статье [6] изучаются возможности искусственных вейвлет-нейронных сетей для прогнозирования краткосрочного суточного стока реки Аджичай в северо-западном регионе Ирана. В работе [7] на основе моделей пошаговой линейной регрессии и географически взвешенной регрессии изучалась взаимосвязь между характером ландшафта и качеством воды в оазисе озера Эбинур в Китае.

Особенностью сибирских рек является то, что они, как правило, берут начало в горных массивах и питаются от ледниковых и снежных вод. Река Ия в Иркутской области берёт начало на северных склонах Восточного Саяна. Длина — 484 км, площадь водосборного бассейна — 18 100 км². слиянием рек Холба и Хиай. Протекает по территории Нижнеудинского, Тулунского, Куйтунского и Братского районов, а также городского округа Тулун. В верховьях имеет горный характер, ниже долина реки расширяется, течение становится спокойным. Внимание к данной реке вызвано крупнейшим за последние годы наводнением, приведшим к жертвам в г. Тулун. Ареалом водосбора реки Ия являются хребты Западного и Восточного

Саяна. Это крупнейшие горные массивы России, протяженностью свыше 600 километров и высотой до 3000 метров. В Восточных Саянах горные вершины нередко называются «белки», а сами хребты — «белогорье». Это отражает то, что эти вершины покрыты снегом круглый год. Климат резко континентальный, с продолжительной и холодной зимой, коротким и прохладным летом. Средняя температура января от $-20\text{ }^{\circ}\text{C}$ (в горах) до $-30\text{ }^{\circ}\text{C}$ (в межгорных котловинах). Средняя температура июля от $+10\text{—}12\text{ }^{\circ}\text{C}$ до $+20\text{ }^{\circ}\text{C}$ соответственно. Количество осадков — 300—350 мм в год в межгорных котловинах, 400—500 мм в северных предгорьях и на южных склонах гор и 1 000—1 200 мм на северных склонах гор. Многие снежники сохраняются в течение всего лета. Питание реки в основном снеговое и дождевое. Весеннее половодье, связанное с таянием снегов, заканчивается в первой половине июня. Летние паводки вызываются длительными дождями. Определенную роль в питании реки играют и грунтовые воды, особенно в летний период.

Обоснование модели

В работе [8] авторами представлена регрессионная модель уровня воды в реке Ия, разработанная на основе статистических данных по выделенным ежемесячным показателям за 2020-2023 гг. Эта модель имеет вид:

$$\begin{aligned} y = & -0.86x_t^1 - 6.96x_t^2 - 8.99x_{t-1}^6 - 2.43x_{t-2}^3 - 1.63x_{t-2}^6 + 7.75x_{t-3}^5 \\ & (-0.36) \quad (-2.96) \quad (-3.49) \quad (-1.05) \quad (-0.67) \quad (4.97) \\ & -8.57x_{t-4}^2 + 25.28x_{t-4}^3 + 6.08x_{t-5}^5 + 13.38x_{t-5}^6 - 2465.01 \quad (1) \\ & (-1.35) \quad (2.92) \quad (5.21) \quad (4.55) \\ & R^2 = 0.89, F = 19.66, E = 7.8\% \end{aligned}$$

Здесь:

y_t – уровень воды в реке Ия в t -ом месяце;

x^1 – количество дней с осадками в горах Саян;

x^2 – средняя дневная температура за месяц, $^{\circ}\text{C}$;

x^3 – средняя ночная температура за месяц, °С;

x^4 – количество осадков в мм за месяц;

x^5 – глубина снежного покрова, см.;

x^6 – среднее атмосферное давление за месяц, мм рт. ст.;

R^2 – критерий множественной детерминации;

F – критерий Фишера;

E – средняя относительная ошибка аппроксимации.

Под каждым коэффициентом модели приведено значение t-критерия Стьюдента.

Таким образом, каждый параметр x имеет два индекса, где первый (верхний) – наименование природного параметра, а нижний (t-τ) указывает на задержку (лаг) в месяцах по отношению к данному месяцу, равную τ.

Отметим, что фактор x^4 не удалось ввести в модель ввиду его статистической незначимости, и это обстоятельство несколько снижает содержательное соответствие модели (1) реальности.

В настоящей работе предпринята попытка повышения адекватности модели (1) путем введения в ее правую часть переменной x^4 , а также фактора сезонности z . Целесообразность дополнительного введения в модель данного фактора вызвана тем, что периоды половодий связаны с сезоном таяния снега и оттаивания промерзшего грунта. В предыдущей модели это не нашло должного внимания и, как результат, именно на этот период – с апреля по июнь приходится наибольшая вариабельность результатов.

Уточненная модель выглядит следующим образом:

$$\begin{aligned} y = & 17.36x_t^2 - 24.61x_t^3 + 1.25x_t^4 + 9.83x_{t-1}^2 + 0.63x_{t-1}^4 - 11.26x_{t-2}^3 \\ & (3.04) \quad (-3.51) \quad (5.73) \quad (3.54) \quad (3.16) \quad (-3.84) \\ & + 5.09x_{t-3}^5 - 5.14x_{t-4}^6 + 7.35x_{t-5}^2 + 3.64x_{t-5}^5 + 39.02z_t + 5130.02 \quad (2) \\ & (4.84) \quad (-2.83) \quad (3.89) \quad (3.29) \quad (4.31) \\ & R^2 = 0.94, F = 33.59, E = 4.8\% \end{aligned}$$

После проведения серии вычислительных экспериментов путем организации конкурса моделей (см., например [9, 10]) сезонная переменная была сформирована в виде, приведенном в табл. 1.

Таблица 1.

Значения сезонной переменной по месяцам

Месяц	Значение z
Январь	1
Февраль	1
Март	1
Апрель	1
Май	3
Июнь	5
Июль	5
Август	5
Сентябрь	5
Октябрь	3
Ноябрь	1
Декабрь	2

Вывод

Сопоставление значений критериев адекватности моделей (1) и (2) указывает на существенно более высокое качество второй из них.

На представленном ниже графике приведены кривые, отражающие реальные значения уровня воды в реке Ия и значения, полученные в результате моделирования. Очевидно, что уточненная модель существенно лучше предыдущей.

Предложенная модель может использоваться для любых рек после соответствующих корректив, посредством обработки статистической информации, что будет рассмотрено в следующих статьях.

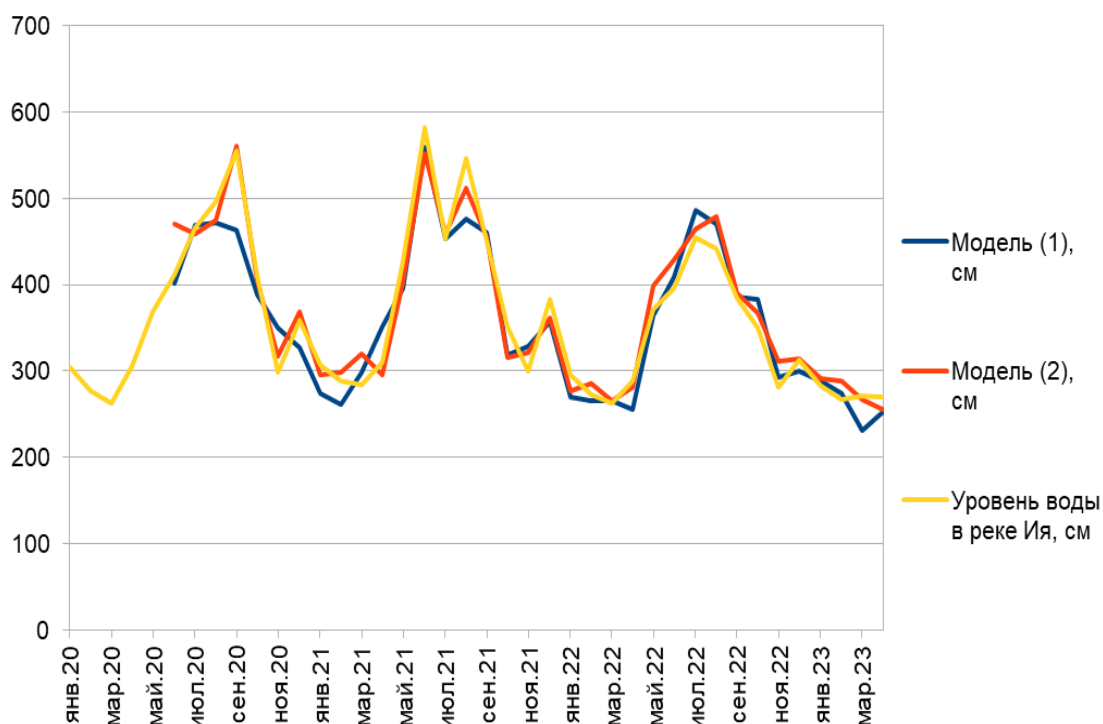


Рис.1. График сезонных изменений уровня воды в реке Ия, населенный пункт Тулун, сравнение моделей (1) и (2) со статистическими значениями

Литература

1. Ajit Pratap Singh, Kunal Dhadse & Jayant Ahalawat Managing water quality of a river using an integrated geographically weighted regression technique with fuzzy decision-making model // *Environmental Monitoring and Assessment*. 2019. №191. URL: doi.org/10.1007/s10661-019-7487-z.

2. Fengyun Sun, Alfonso Mejia & Yue Che Disentangling the Contributions of Climate and Basin Characteristics to Water Yield Across Spatial and Temporal Scales in the Yangtze River Basin. A Combined Hydrological Model and Boosted Regression Approach // *Water Resources Management*. 2019. №33. pp. 3449-3468.

3. Cebrián A. C., Abaurrea J., Asín J. & Segarra E. Dynamic Regression Model for Hourly River Level Forecasting Under Risk Situations: an Application to the Ebro River // *Water Resources Management*. 2019. №33. pp. 523-537.



4. Jang Cheng-Shin Using multi-threshold regression techniques to assess river fecal pollution in the highly urbanized Tamsui River watershed // Environmental Monitoring and Assessment. 2021. №193. URL: doi.org/10.1007/s10661-021-08893-7.

5. Ye Maosheng, Chen Wan, Guo Ling & Li Yuqin “Green” economic development in China: quantile regression evidence from the Yangtze River Economic Belt // Environmental Science and Pollution Research. 2022. №29. pp. 60572–60583.

6. Shafaei Maryam & Kisi Ozgur Predicting river daily flow using wavelet-artificial neural networks based on regression analyses in comparison with artificial neural networks and support vector machine models // Neural Computing and Application. 2017. №28. pp. 15-28.

7. Wang Xiaoping & Zhang Fei Multi-scale analysis of the relationship between landscape patterns and a water quality index (WQI) based on a stepwise linear regression (SLR) and geographically weighted regression (GWR) in the Ebinur Lake oasis // Environmental Science and Pollution Research. 2018. №25. pp. 7033-7048.

8. Носков С.И., Аксенов Ю.Д., Сапожников Ю.М. Регрессионное моделирование уровня воды реки Ия в Иркутской области // Инженерный вестник Дона, 2023, №11. URL: ivdon.ru/ru/magazine/archive/n11y2023/8838

9. Носков С.И., Потороченко Н.А. Диалоговая система реализации конкурса регрессионных зависимостей // Управляющие системы и машины. 1992. №3-4. С. 111-116.

10. Носков С.И. реализация конкурса регрессионных моделей с применением критерия согласованности поведения // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. 2021. №2. С. 153-160.



References

1. Ajit Pratap Singh, Kunal Dhadse & Jayant Ahalawat Environmental Monitoring and Assessment. 2019. №191. URL: doi.org/10.1007/s10661-019-7487-z.
2. Sun Fengyun, Mejia Alfonso & Che Yue Water Resources Management. 2019. №33. pp. 3449-3468.
3. Cebrián A. C., Abaurrea J., Asín J. & Segarra E. Water Resources Management. 2019. №33. pp. 523-537.
4. Jang Cheng-Shin Environmental Monitoring and Assessment. 2021. №193. URL: doi.org/10.1007/s10661-021-08893-7.
5. Ye Maosheng, Chen Wan, Guo Ling & Li Yuqin Environmental Science and Pollution Research. 2022. №29. pp. 60572–60583.
6. Shafaei Maryam & Kisi Ozgur Neural Computing and Application. 2017. №28. pp. 15-28.
7. Wang Xiaoping & Zhang Fei Environmental Science and Pollution Research. 2018. №25. P. 7033-7048.
8. Noskov S.I., Aksenov YU.D., Sapozhnikov YU. M. Inzhenernyj vestnik Dona, 2023, №11. URL: ivdon.ru/ru/magazine/archive/n11y2023/8838 .
9. Noskov S.I., Potorochenko N.A. Upravlyayushchiye sistemy i mashiny. 1992. №3-4. pp. 111-116.
10. Noskov S.I Vestnik Voronezhskogo gosudarstvennogo universiteta. Seriya: Sistemnyj analiz i informatsionnye tekhnologii. 2021. №2. pp. 153-160.

Дата поступления: 21.11.2023

Дата публикации: 6.01.2024